

Normal approximation of Poisson functionals in Kolmogorov distance

Matthias Schulte*
 Institute of Mathematics
 University of Osnabrück, Germany

Abstract

Peccati, Solé, Taqqu and Utzet recently combined Stein's method and Malliavin calculus to obtain a central limit theorem with a bound for the Wasserstein distance of a Poisson functional and a Gaussian random variable. Convergence in the Wasserstein distance always implies convergence in the Kolmogorov distance at a possibly weaker rate. But there are many examples having the same rate for both distances. The aim of this paper is to show this behavior for a large class of Poisson functionals, namely so called U-statistics of Poisson point processes. The technique used by Peccati et al. is modified to establish a similar bound for the Kolmogorov distance of a Poisson functional and a Gaussian random variable. This bound is evaluated for a U-statistic and it is shown that the resulting expression is up to a constant the same as it is for the Wasserstein distance.

Key words: Central limit theorem, Poisson point process, Malliavin calculus, Stein's method, Wiener-Itô chaos expansion

MSC (2010): Primary: 60F05; 60H07 Secondary: 60G55

1 Introduction and results

Let η be a Poisson point process over a Borel space $(X, \mathcal{B}(X))$ with a σ -finite nonatomic intensity measure μ and let $F = F(\eta)$ be a random variable depending on the Poisson point process η . In the following, we call such random variables Poisson functionals. Moreover, we assume that F is square integrable (we write $F \in L^2(\mathbb{P})$) and satisfies $\mathbb{E}F = 0$. By N we denote a standard Gaussian random variable. In [8], Peccati, Solé, Taqqu and Utzet derived by a combination of Stein's method and Malliavin calculus the upper bound

$$d_W(F, N) \leq \mathbb{E}|1 - \langle D_z F, -D_z L^{-1} F \rangle| + \mathbb{E} \int_X (D_z F)^2 |D_z L^{-1} F| \mu(dz) \quad (1)$$

*e-mail: [matthias.schulte\[at\]uni-osnabrueck.de](mailto:matthias.schulte[at]uni-osnabrueck.de)

for the Wasserstein distance of F and N . Here, $\langle \cdot, \cdot \rangle$ stands for the inner product in $L^2(X)$ and the difference operator D and the inverse of the Ornstein-Uhlenbeck generator L are operators from Malliavin calculus. The underlying idea of these operators is that each square integrable Poisson functional has a representation

$$F = \mathbb{E}F + \sum_{n=1}^{\infty} I_n(f_n),$$

where the f_n are square integrable functions, I_n stands for the n -th multiple Wiener-Itô integral and the right hand side converges in $L^2(\mathbb{P})$. This decomposition is called Wiener-Itô chaos expansion and the Malliavin operators of F are defined via their chaos expansion. The operators $D_z F$ and $D_z L^{-1} F$ that occur in (1) are given by

$$D_z F = \sum_{n=1}^{\infty} n I_{n-1}(f_n(z, \cdot)) \quad \text{and} \quad D_z L^{-1} F = - \sum_{n=1}^{\infty} I_{n-1}(f_n(z, \cdot)) \quad \text{for } z \in X.$$

Here, $f_n(z, \cdot)$ stands for the function on X^{n-1} we obtain by taking z as first argument. For exact definitions including the domains and more details on the Malliavin operators we refer to Section 2.

The Wasserstein distance between two random variables Y and Z is defined by

$$d_W(Y, Z) = \sup_{h \in \text{Lip}(1)} |\mathbb{E}h(Y) - \mathbb{E}h(Z)|,$$

where $\text{Lip}(1)$ is the set of all functions $h : \mathbb{R} \rightarrow \mathbb{R}$ with Lipschitz constant less or equal than one. Another commonly used distance for random variables is the Kolmogorov distance

$$d_K(Y, Z) = \sup_{t \in \mathbb{R}} |\mathbb{P}(Y \leq t) - \mathbb{P}(Z \leq t)|,$$

which is the supremum norm of the difference of the distribution functions of Y and Z . Because of this straightforward interpretation, one is often more interested in the Kolmogorov distance than in the Wasserstein distance. For the important case that Z is a standard Gaussian random variable N it is known (see [2, Theorem 3.1]) that

$$d_K(Y, N) \leq 2\sqrt{d_W(Y, N)}. \tag{2}$$

This inequality gives us for the Kolmogorov distance a weaker rate of convergence than for the Wasserstein distance. But for many classical central limit theorems, one has actually the same rate of convergence for both metrics.

In order to overcome the problem that a detour around the Wasserstein distance and the inequality (2) often gives a suboptimal rate of convergence for the Kolmogorov distance, we derive a similar bound as (1) for the Kolmogorov distance by a modification of the proof in [8].

Theorem 1.1 *Let $F \in L^2(\mathbb{P})$ with $\mathbb{E}F = 0$ be in the domain of D and let N be a standard Gaussian random variable. Then*

$$\begin{aligned} d_K(F, N) &\leq \mathbb{E}|1 - \langle D_z F, -D_z L^{-1} F \rangle| + 2\mathbb{E}\langle (D_z F)^2, |D_z L^{-1} F| \rangle \\ &\quad + 2\mathbb{E}\langle (D_z F)^2, |F D_z L^{-1} F| \rangle + 2\mathbb{E}\langle (D_z F)^2, |D_z F D_z L^{-1} F| \rangle \\ &\quad + \sup_{t \in \mathbb{R}} \mathbb{E}\langle D_z \mathbb{1}_{\{F > t\}}, D_z F | D_z L^{-1} F| \rangle \\ &\leq \mathbb{E}|1 - \langle D_z F, -D_z L^{-1} F \rangle| + 2c(F) \sqrt{\mathbb{E}\langle (D_z F)^2, (D_z L^{-1} F)^2 \rangle} \\ &\quad + \sup_{t \in \mathbb{R}} \mathbb{E}\langle D_z \mathbb{1}_{\{F > t\}}, D_z F | D_z L^{-1} F| \rangle \end{aligned} \quad (3)$$

with

$$c(F) = \sqrt{\mathbb{E}\langle (D_z F)^2, (D_z F)^2 \rangle} + (\mathbb{E}\langle D_z F, D_z F \rangle^2)^{\frac{1}{4}} \left((\mathbb{E}F^4)^{\frac{1}{4}} + 1 \right).$$

Comparing (1) and (3), one notes that both terms of the Wasserstein bound (1) also occur in (3), which means that the bound for the Kolmogorov distance is always larger.

We apply our Theorem 1.1 to two situations, where we obtain the same rate of convergence for the Kolmogorov distance and the Wasserstein distance. At first we derive the classical Berry-Esseen inequality with the optimal rate of convergence for the normal approximation of a classical Poisson random variable. As another application of Theorem 1.1, we consider so called U-statistics of Poisson point processes, which are defined as

$$F = \sum_{(x_1, \dots, x_k) \in \eta_{\neq}^k} f(x_1, \dots, x_k)$$

with $k \in \mathbb{N}$, $\eta_{\neq}^k = \{(x_1, \dots, x_k) \in \eta^k : x_i \neq x_j \ \forall i \neq j\}$ and $f \in L^1(X^k)$. In [3, 4] and [10], Lachièze-Rey and Peccati and Reitzner and Schulte used the bound (1) for the Wasserstein distance to derive central limit theorems with explicit rates of convergence for such functionals occurring in stochastic geometry and random graph theory. Now Theorem 1.1 allows us to replace the Wasserstein distance by the Kolmogorov distance without changing the rate of convergence, which means that the inequality (2) is not sharp for this class of functionals.

These applications are discussed in Section 4 and the result for U-statistics is shown in Section 5. Before we prove our main result Theorem 1.1 in Section 3, we introduce some facts from Malliavin calculus and Stein's method in Section 2.

In this paper, we use the following notation. By $L^p(\mathbb{P})$ we denote the set of random variables such that $\mathbb{E}|X|^p < \infty$ and $L^p(X^n)$ stands for the set of functions $f : X^n \rightarrow \overline{\mathbb{R}} := \mathbb{R} \cup \{\pm\infty\}$ satisfying $\int_{X^n} |f|^p d\mu^n = \int_{X^n} |f(x_1, \dots, x_n)|^p \mu(dx_1) \dots \mu(dx_n) < \infty$. The norm in $L^2(X^n)$ is denoted by $\|\cdot\|_n$ and $\langle \cdot, \cdot \rangle_n$ is the inner product in $L^2(X^n)$. For $n = 1$ we write $\langle \cdot, \cdot \rangle$ instead of $\langle \cdot, \cdot \rangle_1$. By $L_s^p(X^n)$ we denote the set of all functions $f \in L^p(X^n)$ that are invariant under permutations of the arguments. The support of a function $f : X^n \rightarrow \overline{\mathbb{R}}$ is defined by $\text{supp}(f) = \{(x_1, \dots, x_n) \in X^n : f(x_1, \dots, x_n) \neq 0\}$.

2 Preliminaries

Malliavin calculus for Poisson functionals. In the sequel, we briefly introduce three Malliavin operators and some properties of them that are necessary for the proofs

in this paper. For more details on Malliavin calculus for Poisson functionals we refer to [5, 7, 8] and the references therein.

By $I_n(\cdot)$, $n \geq 1$, we denote the n -th multiple Wiener-Itô integral, which is defined for all functions $f \in L_s^2(X^n)$ and satisfies $\mathbb{E}I_n(f) = 0$. The multiple Wiener-Itô integrals are orthogonal in the sense that

$$\mathbb{E}I_m(f)I_n(g) = \begin{cases} n!\langle f, g \rangle_n, & m = n \\ 0, & m \neq n \end{cases} \quad (4)$$

for all $f \in L_s^2(X^m)$, $g \in L_s^2(X^n)$, $m, n \geq 1$. It is known (see [5] for a proof) that every Poisson functional $F \in L^2(\mathbb{P})$ has a unique so called **Wiener-Itô chaos expansion**

$$F = \mathbb{E}F + \sum_{n=1}^{\infty} I_n(f_n) \quad (5)$$

with $f_n \in L_s^2(X^n)$, where the series converges in $L^2(\mathbb{P})$. In the following, we call the functions f_n kernels and say that F has a chaos expansion of order k if $f_n = 0$ for all $n > k$. Combining (4) and (5), we obtain

$$\text{Var } F = \sum_{n=1}^{\infty} n! \|f_n\|_n^2.$$

The representation (5) allows us to define the difference operator D , the Ornstein-Uhlenbeck generator L and the Skorohod integral δ in the following way:

Definition 2.1 Let $F \in L^2(\mathbb{P})$ with the Wiener-Itô chaos expansion (5). If $\sum_{n=1}^{\infty} nn! \|f_n\|_n^2 < \infty$, then the random function $z \mapsto D_z F$ defined by

$$D_z F = \sum_{n=1}^{\infty} n I_{n-1}(f_n(z, \cdot)), \quad z \in X$$

is called the **difference operator** of F . For $\sum_{n=1}^{\infty} n^2 n! \|f_n\|_n^2 < \infty$ the **Ornstein-Uhlenbeck generator** of F , denoted by LF , is given by

$$LF = - \sum_{n=1}^{\infty} n I_n(f_n).$$

Let $z \mapsto g(z)$ be a random function with a chaos expansion

$$g(z) = g_0(z) + \sum_{n=1}^{\infty} I_n(g_n(z, \cdot)), \quad g_n(z, \cdot) \in L_s^2(X^n)$$

for every $z \in X$ and $\sum_{n=0}^{\infty} (n+1)! \|g_n\|_{n+1}^2 < \infty$. Then the **Skorohod integral** of g is the random variable $\delta(g)$ defined by

$$\delta(g) = \sum_{n=0}^{\infty} I_{n+1}(\tilde{g}_n),$$

where \tilde{g}_n is the symmetrization $\tilde{g}_n(x_1, \dots, x_{n+1}) = \frac{1}{(n+1)!} \sum_{\sigma} g_n(x_{\sigma(1)}, \dots, x_{\sigma(n+1)})$ over all permutations σ of the $n+1$ variables.

We denote the domains of these operators by $\text{dom } D$, $\text{dom } L$ and $\text{dom } \delta$. The difference operator also has the geometric interpretation

$$D_z F(\eta) = F(\eta + \delta_z) - F(\eta), \quad (6)$$

where δ_z is the Dirac measure concentrated at the point $z \in X$, whence it is sometimes called add-one-cost operator. For centered random variables $F \in L^2(\mathbb{P})$, i.e. $\mathbb{E}F = 0$, the inverse Ornstein-Uhlenbeck generator is given by

$$L^{-1}F = -\sum_{n=1}^{\infty} \frac{1}{n} I_n(f_n).$$

The following lemma summarizes how the operators from Malliavin calculus are related:

Lemma 2.2 *a) For every $F \in \text{dom } L$ it holds that $F \in \text{dom } D$, $DF \in \text{dom } \delta$ and*

$$\delta DF = -LF. \quad (7)$$

b) Let $F \in \text{dom } D$ and $g \in \text{dom } \delta$. Then

$$\mathbb{E}\langle D_z F, g(z) \rangle = \mathbb{E}[F \delta(g)]. \quad (8)$$

For proofs we refer to [8] and [7], respectively. Equation (8) is sometimes called **integration by parts formula**. Because of this identity, one can see the difference operator and the Skorohod integral as dual operators.

For our applications in Section 4 we need the following inequality:

Lemma 2.3 *For $g(z) = I_k(f(z, \cdot))$ with $f(z, \cdot) \in L_s^2(X^k)$ for all $z \in X$, we have*

$$\mathbb{E} [\delta(g)^2] \leq (k+1) \mathbb{E} \int_X I_k(f(z, \cdot))^2 \mu(dz).$$

Proof: By the definition of δ , we obtain $\delta(g) = I_{k+1}(\tilde{f})$ with the symmetrization

$$\tilde{f}(x_1, \dots, x_{k+1}) = \frac{1}{(k+1)!} \sum_{\sigma} f(x_{\sigma(1)}, \dots, x_{\sigma(k+1)})$$

as above. From the Cauchy-Schwarz inequality, it follows $\|\tilde{f}\|_{k+1}^2 \leq \|f\|_{k+1}^2$. Combining this with Fubini's Theorem, we have

$$\mathbb{E} [\delta(g)^2] = (k+1)! \|\tilde{f}\|_{k+1}^2 \leq (k+1)! \|f\|_{k+1}^2 = (k+1) \mathbb{E} \int_X I_k(f(z, \cdot))^2 \mu(dz).$$

□

Stein's method. Besides Malliavin calculus our proof of Theorem 1.1 rests upon Stein's method, which goes back to Charles Stein [11, 12] and is a powerful tool for proving limit theorems. For a detailed and more general introduction into this topic we refer to [1, 2, 12]. Very fundamental for this approach is the following lemma (see Chapter II in [12]):

Lemma 2.4 *The function*

$$g_t(w) = e^{\frac{w^2}{2}} \int_{-\infty}^w (\mathbb{1}_{(-\infty, t]}(s) - \mathbb{P}(N \leq t)) e^{-\frac{s^2}{2}} ds \quad (9)$$

is a solution of the differential equation

$$g'_t(w) - wg_t(w) = \mathbb{1}_{(-\infty, t]}(w) - \mathbb{P}(N \leq t) \quad (10)$$

and satisfies

$$0 < g_t(w) \leq \frac{\sqrt{2\pi}}{4}, \quad |g'_t(w)| \leq 1 \quad \text{and} \quad |wg_t(w)| \leq 1 \quad (11)$$

for any $w \in \mathbb{R}$.

Equation (10) is usually called **Stein's equation**. The function g_t is infinitely differentiable on $\mathbb{R} \setminus \{t\}$, but it is not differentiable in t . We denote the left-sided and right-sided limits of the derivatives in t by $g_t^{(m)}(t-)$ and $g_t^{(m)}(t+)$, respectively. For the first derivative, a direct computation proves

$$g'_t(t+) = -1 + g'_t(t-) \quad (12)$$

and we define $g'_t(t) := g'_t(t-)$.

By replacing w by a random variable Z and taking the expectation in (10), one obtains

$$\mathbb{E}[g'_t(Z) - Zg_t(Z)] = \mathbb{P}(Z \leq t) - \mathbb{P}(N \leq t)$$

and as a consequence of the definition of the Kolmogorov distance

$$d_K(Z, N) = \sup_{t \in \mathbb{R}} |\mathbb{E}[g'_t(Z) - Zg_t(Z)]|. \quad (13)$$

The identity (13) will be our starting point in Section 3. Note furthermore, that we obtain, by combining (10) and (11), the upper bound

$$|g''_t(w)| \leq \frac{\sqrt{2\pi}}{4} + |w|. \quad (14)$$

for $w \in \mathbb{R} \setminus \{t\}$.

3 Proof of Theorem 1.1

By a combination of Malliavin calculus and Stein's method similar to that in [8], we derive the upper bound for the Kolmogorov distance.

Proof of Theorem 1.1: Using the identity (7) and the integration by parts formula (8), we obtain

$$\mathbb{E}[Fg_t(F)] = \mathbb{E}[LL^{-1}Fg_t(F)] = \mathbb{E}[\delta(-D_zL^{-1}F)g_t(F)] = \mathbb{E}\langle -D_zL^{-1}F, D_zg_t(F) \rangle. \quad (15)$$

In order to compute $D_zg_t(F)$, we fix $z \in X$ and consider the following cases:

1. $F, F + D_zF \leq t$ or $F, F + D_zF > t$;
2. $F \leq t < F + D_zF$;
3. $F + D_zF \leq t < F$.

For $F, F + D_zF \leq t$ or $F, F + D_zF > t$, it follows by Taylor expansion that

$$\begin{aligned} D_zg_t(F) &= g_t(F + D_zF) - g_t(F) = g'_t(F)D_zF + \frac{1}{2}g''_t(\tilde{F})(D_zF)^2 \\ &=: g'_t(F)D_zF + r_1(F, z, t), \end{aligned}$$

where \tilde{F} is between F and $F + D_zF$. For $F \leq t < F + D_zF$, we obtain by Taylor expansion and (12)

$$\begin{aligned} D_zg_t(F) &= g_t(F + D_zF) - g_t(F) = g_t(F + D_zF) - g_t(t) + g_t(t) - g_t(F) \\ &= g'_t(t+)(F + D_zF - t) + \frac{1}{2}g''_t(\tilde{F}_1)(F + D_zF - t)^2 \\ &\quad + g'_t(F)(t - F) + \frac{1}{2}g''_t(\tilde{F}_2)(t - F)^2 \\ &= g'_t(F)D_zF + (g'_t(t-) - 1 - g'_t(F))(F + D_zF - t) \\ &\quad + \frac{1}{2}g''_t(\tilde{F}_1)(F + D_zF - t)^2 + \frac{1}{2}g''_t(\tilde{F}_2)(t - F)^2 \\ &= g'_t(F)D_zF - (F + D_zF - t) + g''_t(\tilde{F}_0)(t - F)(F + D_zF - t) \\ &\quad + \frac{1}{2}g''_t(\tilde{F}_1)(F + D_zF - t)^2 + \frac{1}{2}g''_t(\tilde{F}_2)(t - F)^2 \\ &=: g'_t(F)D_zF - (F + D_zF - t) + r_2(F, z, t) \end{aligned}$$

with $\tilde{F}_0, \tilde{F}_1, \tilde{F}_2 \in (F, F + D_zF)$. For $F + D_zF \leq t < F$, we have analogously

$$\begin{aligned} D_zg_t(F) &= g_t(F + D_zF) - g_t(F) = g_t(F + D_zF) - g_t(t) + g_t(t) - g_t(F) \\ &= g'_t(t-)(F + D_zF - t) + \frac{1}{2}g''_t(\tilde{F}_1)(F + D_zF - t)^2 \\ &\quad + g'_t(F)(t - F) + \frac{1}{2}g''_t(\tilde{F}_2)(t - F)^2 \\ &= g'_t(F)D_zF + (g'_t(t+) + 1 - g'_t(F))(F + D_zF - t) \\ &\quad + \frac{1}{2}g''_t(\tilde{F}_1)(F + D_zF - t)^2 + \frac{1}{2}g''_t(\tilde{F}_2)(t - F)^2 \\ &= g'_t(F)D_zF + (F + D_zF - t) + g''_t(\tilde{F}_0)(t - F)(F + D_zF - t) \\ &\quad + \frac{1}{2}g''_t(\tilde{F}_1)(F + D_zF - t)^2 + \frac{1}{2}g''_t(\tilde{F}_2)(t - F)^2 \\ &=: g'_t(F)D_zF + (F + D_zF - t) + r_2(F, z, t) \end{aligned}$$

with $\tilde{F}_0, \tilde{F}_1, \tilde{F}_2 \in (F + D_z F, F)$. Thus, $D_z g_t(F)$ has a representation

$$D_z g_t(F) = g'_t(F) D_z F + R(F, z, t), \quad (16)$$

where $R(F, z, t)$ is given by

$$\begin{aligned} R(F, z, t) &= \left(\mathbb{1}_{\{F, F+D_z F \leq t\}} + \mathbb{1}_{\{F, F+D_z F > t\}} \right) r_1(F, z, t) \\ &\quad + \left(\mathbb{1}_{\{F \leq t < F+D_z F\}} + \mathbb{1}_{\{F+D_z F \leq t < F\}} \right) (r_2(F, z, t) - |F + D_z F - t|). \end{aligned}$$

Combining (15) and (16) yields

$$\mathbb{E} [g'_t(F) - F g_t(F)] = \mathbb{E} [g'_t(F) - \langle g'_t(F) D_z F + R(F, z, t), -D_z L^{-1} F \rangle]$$

and the triangle inequality and $|g'_t(F)| \leq 1$ lead to

$$\begin{aligned} |\mathbb{E} [g'_t(F) - F g_t(F)]| &\leq |\mathbb{E} [g'_t(F) (1 - \langle D_z F, -D_z L^{-1} F \rangle)]| \\ &\quad + |\mathbb{E} \langle R(F, z, t), D_z L^{-1} F \rangle| \\ &\leq \mathbb{E} |1 - \langle D_z F, -D_z L^{-1} F \rangle| + \mathbb{E} |\langle R(F, z, t), D_z L^{-1} F \rangle|. \end{aligned} \quad (17)$$

In $r_2(F, z, t)$, we assume that t is between F and $F + D_z F$, so that

$$|F + D_z F - t| \leq |D_z F| \quad \text{and} \quad |F - t| \leq |D_z F|.$$

The inequality (14) allows us to bound all second derivatives in $R(F, z, t)$ by

$$|g''_t(\tilde{F}_i)| \leq \frac{\sqrt{2\pi}}{4} + |F| + |D_z F|.$$

Now it is easy to see that

$$\begin{aligned} |R(F, z, t)| &\leq \left(\mathbb{1}_{\{F, F+D_z F \leq t\}} + \mathbb{1}_{\{F, F+D_z F > t\}} \right) \frac{1}{2} \left(\frac{\sqrt{2\pi}}{4} + |F| + |D_z F| \right) (D_z F)^2 \\ &\quad + \left(\mathbb{1}_{\{F \leq t < F+D_z F\}} + \mathbb{1}_{\{F+D_z F \leq t < F\}} \right) |D_z F| \\ &\quad + \left(\mathbb{1}_{\{F \leq t < F+D_z F\}} + \mathbb{1}_{\{F+D_z F \leq t < F\}} \right) 2 \left(\frac{\sqrt{2\pi}}{4} + |F| + |D_z F| \right) (D_z F)^2 \\ &\leq 2 \left(\frac{\sqrt{2\pi}}{4} + |F| + |D_z F| \right) (D_z F)^2 \\ &\quad + \left(\mathbb{1}_{\{F \leq t < F+D_z F\}} + \mathbb{1}_{\{F+D_z F \leq t < F\}} \right) |D_z F|, \end{aligned}$$

where the last summand can be rewritten as

$$\left(\mathbb{1}_{\{F \leq t < F+D_z F\}} + \mathbb{1}_{\{F+D_z F \leq t < F\}} \right) |D_z F| = D_z \mathbb{1}_{\{F > t\}} D_z F.$$

Hence, it follows directly

$$\begin{aligned} \mathbb{E} |\langle R(F, z, t), D_z L^{-1} F \rangle| &\leq 2\mathbb{E} \langle (D_z F)^2, |D_z L^{-1} F| \rangle + 2\mathbb{E} \langle (D_z F)^2, |F D_z L^{-1} F| \rangle \\ &\quad + 2\mathbb{E} \langle (D_z F)^2, |D_z F D_z L^{-1} F| \rangle \\ &\quad + \mathbb{E} \langle D_z \mathbb{1}_{\{F > t\}} D_z F, |D_z L^{-1} F| \rangle \end{aligned}$$

and putting this in (17) concludes the proof of the first inequality in (3). The second bound in (3) is a direct consequence of the Cauchy-Schwarz inequality. \square

In [8], the right hand side of (15) is evaluated for twice differentiable functions $f : \mathbb{R} \rightarrow \mathbb{R}$ with $\sup_{x \in \mathbb{R}} |f'(x)| \leq 1$ and $\sup_{x \in \mathbb{R}} |f''(x)| \leq 2$ (these are the test functions for the Wasserstein distance) instead of the functions g_t as defined in (9). For such a function f it holds that

$$D_z f(F) = f'(F) D_z F + \tilde{r}(F)$$

with $|\tilde{r}(F)| \leq (D_z F)^2$. Since this representation is easier than the representation we obtain for $D_z g_t(F)$, the bound for the Wasserstein distance in (1) is shorter and easier to evaluate than the bound for the Kolmogorov distance in (3).

The supremum-expression in (3) can be further bounded in the following way. By the integration by parts formula (8) and the Cauchy-Schwarz inequality, we have

$$\begin{aligned} \mathbb{E} \langle D_z \mathbb{1}_{\{F > t\}}, D_z F | D_z L^{-1} F \rangle &= \mathbb{E} [\mathbb{1}_{\{F > t\}} \delta(D_z F | D_z L^{-1} F)] \\ &\leq \mathbb{E} \left[\delta(D_z F | D_z L^{-1} F)^2 \right]^{\frac{1}{2}}, \end{aligned}$$

which yields the upper bound

$$\sup_{t \in \mathbb{R}} \mathbb{E} \langle D_z \mathbb{1}_{\{F > t\}}, D_z F | D_z L^{-1} F \rangle \leq \mathbb{E} \left[\delta(D_z F | D_z L^{-1} F)^2 \right]^{\frac{1}{2}}. \quad (18)$$

4 Applications of Theorem 1.1

Normal approximation of a Poisson random variable. As a first application of Theorem 1.1, we compute the Kolmogorov distance between a Poisson random variable Y with parameter $\lambda > 0$ and a normal distribution. In Example 3.5 in [8], the bound (1) is used to compute the Wasserstein distance and the known optimal rate of convergence $\lambda^{-\frac{1}{2}}$ is obtained.

Y has the same distribution as $F = |\eta| = \sum_{x \in \eta} 1$, where η is a Poisson point process on $[0, 1]$ with λ times the Lebesgue measure as intensity measure. The representation

$$I_1(f) = \sum_{x \in \eta} f(x) - \int_X f(x) \mu(dx)$$

for a Wiener-Itô integral of a function $f \in L^1(X) \cap L^2(X)$ and the fact that

$$F = \lambda \int_0^1 1 \, dx + \sum_{x \in \eta} 1 - \lambda \int_0^1 1 \, dx$$

imply that F has the Wiener-Itô chaos expansion $F = \mathbb{E}F + I_1(f_1) = \lambda + I_1(1)$. Hence, the normalized random variable

$$G = \frac{F - \mathbb{E}F}{\sqrt{\text{Var } F}} = \frac{F - \lambda}{\sqrt{\lambda}}$$

has the chaos expansion $G = \frac{1}{\sqrt{\lambda}} I_1(1)$ and $D_z G = -D_z L^{-1} G = \frac{1}{\sqrt{\lambda}}$ for $z \in [0, 1]$. It is easy to see that

$$\mathbb{E}|1 - \langle D_z G, -D_z L^{-1} G \rangle| = |1 - \frac{1}{\lambda} \langle 1, 1 \rangle| = |1 - \frac{\lambda}{\lambda}| = 0$$

and we obtain

$$\mathbb{E}\langle (D_z G)^2, (D_z L^{-1} G)^2 \rangle = \mathbb{E}\langle (D_z G)^2, (D_z G)^2 \rangle = \frac{1}{\lambda},$$

$\mathbb{E}\langle D_z G, D_z G \rangle^2 = 1$ and $\mathbb{E}G^4 = 3 + \frac{1}{\lambda}$ by analogous computations. Moreover, it holds that

$$\mathbb{E}[\delta(D_z G | D_z L^{-1} G)^2] = \frac{1}{\lambda^2} \mathbb{E}[\delta(1)^2] = \frac{1}{\lambda^2} \mathbb{E}[I_1(1)^2] = \frac{1}{\lambda},$$

so that Theorem 1.1 combined with (18) yields

$$d_K \left(\frac{Y - \lambda}{\sqrt{\lambda}}, N \right) \leq 2 \left(\frac{1}{\sqrt{\lambda}} + \left(3 + \frac{1}{\lambda} \right)^{\frac{1}{4}} + 1 \right) \frac{1}{\sqrt{\lambda}} + \frac{1}{\sqrt{\lambda}} \leq \frac{8}{\sqrt{\lambda}}$$

for $\lambda \geq 1$, which is the classical Berry-Esseen inequality with the optimal rate of convergence (up to a constant).

Normal approximation of U-statistics of Poisson point processes. As a second application of Theorem 1.1, we discuss U-statistics of the form

$$F = \sum_{(x_1, \dots, x_k) \in \eta_{\neq}^k} f(x_1, \dots, x_k)$$

with $k \in \mathbb{N}$, $f \in L_s^1(X^k)$ and $F \in L^2(\mathbb{P})$. Here, η_{\neq}^k stands for the set of all k -tuples of distinct points of η . From now on, we denote k as the order of the U-statistic F . In [10], the chaos expansions of such Poisson functionals are investigated and the bound (1) is used to prove a central limit theorem with a bound for the Wasserstein distance. From there it is known that the kernels of the chaos expansion of a U-statistic $F \in L^2(\mathbb{P})$ are

$$f_i(x_1, \dots, x_i) = \binom{k}{i} \int_{X^{k-i}} f(x_1, \dots, x_i, y_1, \dots, y_{k-i}) \mu(dy_1) \dots \mu(dy_{k-i}) \quad (19)$$

for $i \leq k$ and $f_i = 0$ for $i > k$. An application of the bound (1) to such functionals yields (see Theorem 4.1 in [10])

$$d_W \left(\frac{F - \mathbb{E}F}{\sqrt{\text{Var } F}}, N \right) \leq k \sum_{i,j=1}^k \frac{\sqrt{R_{ij}}}{\sqrt{\text{Var } F}} + k^{\frac{7}{2}} \sum_{i=1}^k \frac{\sqrt{\tilde{R}_i}}{\sqrt{\text{Var } F}}, \quad (20)$$

where R_{ij} and \tilde{R}_i are given by

$$\begin{aligned} R_{ij} &= \mathbb{E} \left(\int_X I_{i-1}(f_i(z, \cdot)) I_{j-1}(f_j(z, \cdot)) \mu(dz) \right)^2 \\ &\quad - \left(\mathbb{E} \int_X I_{i-1}(f_i(z, \cdot)) I_{j-1}(f_j(z, \cdot)) \mu(dz) \right)^2 \\ \tilde{R}_i &= \mathbb{E} \int_X I_{i-1}(f_i(z, \cdot))^4 \mu(dz) \end{aligned}$$

for $i, j = 1, \dots, k$. In [10], the right hand side of (20) is bounded by a sum of deterministic integrals $M_{ij}(f)$ depending on f . Due to technical reasons it is assumed that the U-statistic F is absolutely convergent, which means that the U-statistic \overline{F} given by

$$\overline{F} = \sum_{(x_1, \dots, x_k) \in \eta_{\neq}^k} |f(x_1, \dots, x_k)|$$

is in $L^2(\mathbb{P})$. The U-statistic \overline{F} has a finite Wiener-Itô chaos expansion with kernels

$$\overline{f}_i(x_1, \dots, x_i) = \binom{k}{i} \int_{X^{k-i}} |f(x_1, \dots, x_i, y_1, \dots, y_{k-i})| \mu(dy_1) \dots \mu(dy_{k-i})$$

for $i = 1, \dots, k$ and $\overline{f}_i = 0$ for $i > k$. In order to define $M_{ij}(f)$ as in [10], we use the following notation. For $i, j = 1, \dots, k$ let $\overline{\Pi}_{\geq 2}(i, j, i, j)$ be the set of all partitions π of

$$x_1^{(1)}, \dots, x_i^{(1)}, x_1^{(2)}, \dots, x_j^{(2)}, x_1^{(3)}, \dots, x_i^{(3)}, x_1^{(4)}, \dots, x_j^{(4)}$$

such that

- two variables with the same upper index are in different blocks of π ;
- each block of π includes at least two variables;
- there are no sets $A_1, A_2 \subset \{1, 2, 3, 4\}$ with $A_1 \cup A_2 = \{1, 2, 3, 4\}$ and $A_1 \cap A_2 = \emptyset$ such that each block of π either consists of variables with upper index in A_1 or of variables with upper index in A_2 .

Let $|\pi|$ stand for the number of blocks of a partition π . For functions $g_1, g_3 : X^i \rightarrow \overline{\mathbb{R}}$ and $g_2, g_4 : X^j \rightarrow \overline{\mathbb{R}}$ the tensor product $g_1 \otimes g_2 \otimes g_3 \otimes g_4 : X^{i+j+i+j} \rightarrow \overline{\mathbb{R}}$ is given by

$$\begin{aligned} & (g_1 \otimes g_2 \otimes g_3 \otimes g_4)(x_1^{(1)}, \dots, x_j^{(4)}) \\ &= g_1(x_1^{(1)}, \dots, x_i^{(1)}) g_2(x_1^{(2)}, \dots, x_j^{(2)}) g_3(x_1^{(3)}, \dots, x_i^{(3)}) g_4(x_1^{(4)}, \dots, x_j^{(4)}). \end{aligned}$$

For $\pi \in \overline{\Pi}_{\geq 2}(i, j, i, j)$ we define the function $(g_1 \otimes g_2 \otimes g_3 \otimes g_4)_\pi(x_1^{(1)}, \dots, x_j^{(4)}) : X^{|\pi|} \rightarrow \overline{\mathbb{R}}$ by replacing all variables that are in the same block of π by a new variable. Since we later integrate over all these new variables, their order does not matter. Using this notation, we define

$$M_{ij}(f) = \sum_{\pi \in \overline{\Pi}_{\geq 2}(i, j, i, j)} \int_{X^{|\pi|}} (\overline{f}_i \otimes \overline{f}_j \otimes \overline{f}_i \otimes \overline{f}_j)_\pi d\mu^{|\pi|}$$

for $i, j = 1, \dots, k$. Now we can state the following upper bound for the Wasserstein distance (see Theorem 4.7 in [10]):

Proposition 4.1 *Let $F \in L^2(\mathbb{P})$ be an absolutely convergent U-statistic of order k and let N be a standard Gaussian random variable. Then we have*

$$d_W \left(\frac{F - \mathbb{E}F}{\sqrt{\text{Var } F}}, N \right) \leq 2k^{\frac{7}{2}} \sum_{1 \leq i \leq j \leq k} \frac{\sqrt{M_{ij}(f)}}{\text{Var } F}. \quad (21)$$

In this situation, we can use Theorem 1.1 to prove the following bound analogous to (21) for the Kolmogorov distance between a U-statistic and a Gaussian random variable:

Theorem 4.2 *Let $F \in L^2(\mathbb{P})$ be an absolutely convergent U-statistic of order k and let N be a standard Gaussian random variable. Then we have*

$$d_K \left(\frac{F - \mathbb{E}F}{\sqrt{\text{Var } F}}, N \right) \leq 17k^6 \sum_{i,j=1}^k \frac{\sqrt{M_{ij}(f)}}{\text{Var } F}. \quad (22)$$

Before we prove this theorem in Section 5, we discuss some of its consequences. Let us consider a family of Poisson point processes η_t with intensity measures μ_t and U-statistics

$$F_t = \sum_{(x_1, \dots, x_k) \in (\eta_t)_\neq^k} f_t(x_1, \dots, x_k)$$

with $f_t \in L_s^1(X^k)$ and $F_t \in L^2(\mathbb{P})$ such that

$$\frac{\sqrt{M_{ij}(f_t)}}{\text{Var } F_t} \rightarrow 0 \text{ as } t \rightarrow \infty \text{ for all } i, j = 1, \dots, k.$$

Here, we integrate with respect to μ_t in $L^1(X^k)$ and $M_{ij}(f)$. Comparing the right hand sides in (21) and (22) for the U-statistics F_t , we see that the bounds for the Wasserstein and Kolmogorov distance have the same rates of convergence and differ only by constants. An important special case of the described setting is that the Poisson point process depends on a real valued intensity parameter. The following corollary deals with this situation and is the counterpart of Theorem 5.1 in [10] and Theorem 5.1 in [6] for the Kolmogorov distance.

Corollary 4.3 *Let η_t be a Poisson point process with intensity measure $\mu_t = t\mu$ with $t \geq 1$ and a fixed σ -finite non-atomic measure μ and let N be a standard Gaussian random variable. We consider U-statistics $F_t \in L^2(\mathbb{P})$ of the form*

$$F_t = g(t) \sum_{(x_1, \dots, x_k) \in \eta_{t,\neq}^k} f(x_1, \dots, x_k)$$

with $g : (0, \infty) \rightarrow (0, \infty)$ and $f \in L_s^1(X^k)$ independent of t . Moreover, we assume

$$\int_X \left(\int_{X^{k-1}} f(x, y_1, \dots, y_{k-1}) \mu(dy_1) \dots \mu(dy_{k-1}) \right)^2 \mu(dx) > 0$$

and $M_{ij}(f) < \infty$ for $i, j = 1, \dots, k$. Then there is a constant $C > 0$ such that

$$d_K \left(\frac{F_t - \mathbb{E}F_t}{\sqrt{\text{Var } F_t}}, N \right) \leq Ct^{-\frac{1}{2}}$$

for all $t \geq 1$.

In [3], a so called fourth moment condition for Poisson functionals with finite Wiener-Itô chaos expansion and non-negative kernels satisfying some integrability conditions was derived. More precisely, for such Poisson functionals it was proven that

$$d_W \left(\frac{F - \mathbb{E}F}{\sqrt{\text{Var } F}}, N \right) \leq C_{W,k} \sqrt{\mathbb{E} \left(\frac{F - \mathbb{E}F}{\sqrt{\text{Var } F}} \right)^4 - 3}$$

with a constant $C_{W,k} > 0$ only depending on k . U-statistics $F \in L^2(\mathbb{P})$ of the form

$$F = \sum_{(x_1, \dots, x_k) \in \eta_{\neq}^k} f(x_1, \dots, x_k) \quad \text{with } f \in L_s^1(X^k) \text{ and } f \geq 0$$

such that $M_{ij}(f) < \infty$ for $i, j = 1, \dots, k$ belong to this class and satisfy

$$\frac{M_{ij}(f)}{(\text{Var } F)^2} \leq \mathbb{E} \left(\frac{F - \mathbb{E}F}{\sqrt{\text{Var } F}} \right)^4 - 3.$$

Then (22) can be modified to

$$d_K \left(\frac{F - \mathbb{E}F}{\sqrt{\text{Var } F}}, N \right) \leq C_k \sqrt{\mathbb{E} \left(\frac{F - \mathbb{E}F}{\sqrt{\text{Var } F}} \right)^4 - 3}$$

with a constant $C_k > 0$ only depending on k .

5 Proof of Theorem 4.2

In our proof of Theorem 4.2, we make use of the following property of U-statistics:

Lemma 5.1 *For a U-statistic $F \in L^2(\mathbb{P})$ of order k the inverse of the Ornstein-Uhlenbeck generator has a representation*

$$\begin{aligned} -L^{-1}(F - \mathbb{E}F) &= \sum_{m=1}^k \frac{1}{m} \sum_{(x_1, \dots, x_m) \in \eta_{\neq}^m} \int_{X^{k-m}} f(x_1, \dots, x_m, y_1, \dots, y_{k-m}) \mu(dy_1) \dots \mu(dy_{k-m}) \\ &\quad - \sum_{m=1}^k \frac{1}{m} \int_{X^k} f(y_1, \dots, y_k) \mu(dy_1) \dots \mu(dy_k). \end{aligned} \quad (23)$$

Proof: We define $\widehat{f}_i : X^i \rightarrow \overline{\mathbb{R}}$ by $\widehat{f}_i(x_1, \dots, x_i) = \binom{k}{i}^{-1} f_i(x_1, \dots, x_i)$ for $1 \leq i \leq k$. Using this notation and formula (19) for the kernels of a U-statistic, we obtain the chaos expansion

$$\begin{aligned} &\sum_{(x_1, \dots, x_m) \in \eta_{\neq}^m} \int_{X^{k-m}} f(x_1, \dots, x_m, y_1, \dots, y_{k-m}) \mu(dy_1) \dots \mu(dy_{k-m}) \\ &= \int_{X^k} f(y_1, \dots, y_k) \mu(dy_1) \dots \mu(dy_k) + \sum_{i=1}^m \binom{m}{i} I_i(\widehat{f}_i) \end{aligned}$$

for $1 \leq m \leq k$. Combining this with an identity for binomial coefficients, we see that the right hand side in (23) equals

$$\begin{aligned} \sum_{m=1}^k \frac{1}{m} \sum_{i=1}^m \binom{m}{i} I_i(\widehat{f}_i) &= \sum_{m=1}^k \sum_{i=1}^k \frac{1}{m} \binom{m}{i} I_i(\widehat{f}_i) = \sum_{i=1}^k \sum_{m=1}^k \frac{1}{m} \binom{m}{i} I_i(\widehat{f}_i) \\ &= \sum_{i=1}^k \frac{1}{i} \binom{k}{i} I_i(\widehat{f}_i) = \sum_{i=1}^k \frac{1}{i} I_i(f_i), \end{aligned}$$

which is the chaos expansion of $-L^{-1}(F - \mathbb{E}F)$ by definition. \square

In order to deal with expressions as R_{ij} and \tilde{R}_i in the previous section, one needs to compute the expectation of products of multiple Wiener-Itô integrals. This can be done by using Theorem 3.1 in [6] or similar results from [9, 13]. This so called product formula tells us that the expectation of a product of multiple Wiener-Itô integrals is a sum of integrals of deterministic functions depending on the integrands of the stochastic integrals and partitions of their variables as used for the definition of $M_{ij}(f)$ and requires that some technical integrability conditions are satisfied. This is always the case if the integrands are bounded and have bounded support. In order to ensure that all functions f_i and \tilde{f}_i have this property, we restrict ourselves at first to U-statistics $F \in L^2(\mathbb{P})$ of the form

$$F = \sum_{(x_1, \dots, x_k) \in \eta_{\neq}^k} f(x_1, \dots, x_k) \quad \text{with } f \in L_s^1(X^k) \text{ bounded and } \mu^k(\text{supp}(f)) < \infty. \quad (24)$$

Then, one can prove as in [10], by using the product formula for multiple Wiener-Itô integrals, that

$$R_{ij} \leq M_{ij}(f) \quad \text{and} \quad \tilde{R}_i \leq M_{ii}(f) \quad \text{for } i, j = 1, \dots, k. \quad (25)$$

We also obtain the following bound for the fourth moment:

Lemma 5.2 *For a U-statistic F satisfying (24) it holds that*

$$\mathbb{E}(F - \mathbb{E}F)^4 \leq \left(\sum_{i,j=1}^k \sqrt{M_{ij}(f)} + k^2 \text{Var } F \right)^2.$$

Proof. Using the Wiener-Itô chaos expansion of F and the Cauchy-Schwarz inequality, we have

$$\begin{aligned} \mathbb{E}(F - \mathbb{E}F)^4 &= \sum_{n_1, n_2, n_3, n_4=1}^k \mathbb{E} I_{n_1}(f_{n_1}) I_{n_2}(f_{n_2}) I_{n_3}(f_{n_3}) I_{n_4}(f_{n_4}) \\ &\leq \sum_{n_1, n_2, n_3, n_4=1}^k \sqrt{\mathbb{E} I_{n_1}(f_{n_1})^2 I_{n_2}(f_{n_2})^2} \sqrt{\mathbb{E} I_{n_3}(f_{n_3})^2 I_{n_4}(f_{n_4})^2} \\ &= \left(\sum_{n_1, n_2=1}^k \sqrt{\mathbb{E} I_{n_1}(f_{n_1})^2 I_{n_2}(f_{n_2})^2} \right)^2. \end{aligned}$$

The assumptions on f allow us to apply the product formula for multiple Wiener-Itô integrals, which yields

$$\mathbb{E} I_{n_1}(f_{n_1})^2 I_{n_2}(f_{n_2})^2 \leq M_{n_1 n_2}(f) + n_1! \|f_{n_1}\|^2 n_2! \|f_{n_2}\|^2 \leq M_{n_1 n_2}(f) + (\text{Var } F)^2.$$

□

Lemma 5.3 *Let F be a U-statistic of the form (24). Then we have*

$$\sup_{t \in \mathbb{R}} \mathbb{E} \langle D_z \mathbb{1}_{\{F > t\}}, D_z F | D_z L^{-1}(F - \mathbb{E}F) \rangle \leq \sqrt{(2k-1) \mathbb{E} \langle (D_z \bar{F})^2, (D_z L^{-1}(\bar{F} - \mathbb{E}\bar{F}))^2 \rangle}.$$

Proof: We can write the U-statistic F as

$$F = \sum_{(x_1, \dots, x_k) \in \eta_{\neq}^k} f(x_1, \dots, x_k) = \underbrace{\sum_{(x_1, \dots, x_k) \in \eta_{\neq}^k} f^+(x_1, \dots, x_k)}_{=F^+} - \underbrace{\sum_{(x_1, \dots, x_k) \in \eta_{\neq}^k} f^-(x_1, \dots, x_k)}_{=F^-}$$

with $f^+ = \max\{f, 0\}$ and $f^- = \max\{-f, 0\}$ and $\bar{F} = F^+ + F^-$. As a consequence of (6), we know that $D_z V \geq 0$ for a U-statistic V where we sum over a non-negative function. Combining this with $f^+, f^- \geq 0$ and Lemma 5.1, we see that

$$-D_z L^{-1}(F^+ - \mathbb{E}F^+) \geq 0 \quad \text{and} \quad -D_z L^{-1}(F^- - \mathbb{E}F^-) \geq 0.$$

Moreover, it holds that $D_z \mathbb{1}_{\{F > t\}} D_z F \geq 0$. Together with the integration by parts formula (8) and the Cauchy-Schwarz inequality, we have

$$\begin{aligned} & \sup_{t \in \mathbb{R}} \mathbb{E} \langle D_z \mathbb{1}_{\{F > t\}}, D_z F | D_z L^{-1}(F - \mathbb{E}F) \rangle \\ &= \sup_{t \in \mathbb{R}} \mathbb{E} \langle D_z \mathbb{1}_{\{F > t\}}, D_z F | D_z L^{-1}(F^+ - \mathbb{E}F^+ - F^- + \mathbb{E}F^-) \rangle \\ &\leq \sup_{t \in \mathbb{R}} \mathbb{E} \langle D_z \mathbb{1}_{\{F > t\}}, D_z F (-D_z L^{-1}(F^+ - \mathbb{E}F^+) - D_z L^{-1}(F^- - \mathbb{E}F^-)) \rangle \\ &\leq \mathbb{E} \left[\delta (D_z F D_z L^{-1}(\bar{F} - \mathbb{E}\bar{F}))^2 \right]^{\frac{1}{2}}. \end{aligned}$$

Now Proposition 3.1 in [13] implies that the product $D_z F D_z L^{-1}(\bar{F} - \mathbb{E}\bar{F})$ has a finite chaos expansion with an order less or equal than $2k - 2$. Hence, we can use Lemma 2.3 to obtain

$$\sup_{t \in \mathbb{R}} \mathbb{E} \langle D_z \mathbb{1}_{\{F > t\}}, D_z F | D_z L^{-1} F \rangle \leq \sqrt{(2k-1) \mathbb{E} \langle (D_z F)^2, (D_z L^{-1}(\bar{F} - \mathbb{E}\bar{F}))^2 \rangle}$$

and the fact that $(D_z F)^2 \leq (D_z \bar{F})^2$ concludes the proof. □

Proof of Theorem 4.2: At first, we assume that F has the form (24), which allows us to use the inequalities in (25).

We consider the normalized random variable $G = \frac{F - \mathbb{E}F}{\sqrt{\text{Var } F}}$, whose Wiener-Itô chaos expansion has the kernels $g_i = \frac{1}{\sqrt{\text{Var } F}} f_i$ for $i = 1, \dots, k$. In order to simplify our notation, we use the abbreviation

$$S = \sum_{i,j=1}^k \frac{\sqrt{M_{ij}(f)}}{\text{Var } F}.$$

Exactly as in the proof of Theorem 4.1 in [10], we obtain

$$\mathbb{E} |1 - \langle D_z G, -D_z L^{-1} G \rangle| \leq k \sum_{i,j=1}^k \frac{\sqrt{R_{ij}}}{\text{Var } F} \leq k \sum_{i,j=1}^k \frac{\sqrt{M_{ij}(f)}}{\text{Var } F} = kS.$$

By straightforward computations using Fubini's Theorem and the Cauchy-Schwarz inequality, it follows

$$\begin{aligned} \mathbb{E} \langle (D_z G)^2, (D_z L^{-1} G)^2 \rangle &= \int_X \mathbb{E} [(D_z G)^2 (D_z L^{-1} G)^2] \mu(dz) \\ &= \int_X \mathbb{E} \left[\left(\sum_{i=1}^k i I_{i-1}(g_i(z, \cdot)) \right)^2 \left(\sum_{i=1}^k I_{i-1}(g_i(z, \cdot)) \right)^2 \right] \mu(dz) \\ &\leq \int_X k^5 \sum_{i=1}^k \mathbb{E} [I_{i-1}(g_i(z, \cdot))^4] \mu(dz) \leq k^5 \sum_{i=1}^k \frac{\tilde{R}_i}{(\text{Var } F)^2} \\ &\leq k^5 \sum_{i=1}^k \frac{M_{ii}(f)}{(\text{Var } F)^2} \leq k^5 S^2, \end{aligned}$$

$$\begin{aligned} \mathbb{E} \langle D_z G^2, D_z G^2 \rangle &= \int_X \mathbb{E} \left[\left(\sum_{i=1}^k i I_{i-1}(g_i(z, \cdot)) \right)^4 \right] \mu(dz) \\ &\leq \int_X k^3 \sum_{i=1}^k i^4 \mathbb{E} [I_{i-1}(g_i(z, \cdot))^4] \mu(dz) \leq k^7 \sum_{i=1}^k \frac{\tilde{R}_i}{(\text{Var } F)^2} \\ &\leq k^7 \sum_{i=1}^k \frac{M_{ii}(f)}{(\text{Var } F)^2} \leq k^7 S^2 \end{aligned}$$

and

$$\begin{aligned}
\mathbb{E}\langle D_z G, D_z G \rangle^2 &\leq \mathbb{E} \left(\sum_{i,j=1}^k i j \int_X I_{i-1}(g_i(z, \cdot)) I_{j-1}(g_j(z, \cdot)) \mu(dz) \right)^2 \\
&\leq k^2 \sum_{i,j=1}^k i^2 j^2 \mathbb{E} \left(\int_X I_{i-1}(g_i(z, \cdot)) I_{j-1}(g_j(z, \cdot)) \mu(dz) \right)^2 \\
&\leq k^6 \sum_{i,j=1}^k \frac{R_{ij}}{(\text{Var } F)^2} + k^4 \sum_{i=1}^k \frac{(i!)^2 \|f_i\|_i^4}{(\text{Var } F)^2} \\
&\leq k^6 \sum_{i,j=1}^k \frac{R_{ij}}{(\text{Var } F)^2} + k^4 \leq k^6 \sum_{i,j=1}^k \frac{M_{ij}(f)}{(\text{Var } F)^2} + k^4 \leq k^6 S^2 + k^4.
\end{aligned}$$

As a consequence of Lemma 5.2, we have that

$$\mathbb{E} \left(\frac{F - \mathbb{E}F}{\sqrt{\text{Var } F}} \right)^4 \leq \left(\sum_{i,j=1}^k \frac{\sqrt{M_{ij}(f)}}{\text{Var } F} + k^2 \right)^2 = (S + k^2)^2.$$

Lemma 5.3 combined with a similar computation as for $\mathbb{E}\langle (D_z G)^2, (D_z L^{-1} G)^2 \rangle$ implies that

$$\sup_{t \in \mathbb{R}} \mathbb{E}\langle D_z \mathbb{1}_{G>t}, D_z G | D_z L^{-1} G \rangle \leq \sqrt{(2k-1)k^5 \sum_{i=1}^k \frac{M_{ii}(f)}{(\text{Var } F)^2}} \leq \sqrt{2}k^3 S.$$

Thus, it follows from Theorem 1.1 that

$$\begin{aligned}
d_K \left(\frac{F - \mathbb{E}F}{\sqrt{\text{Var } F}}, N \right) &\leq kS + 2(k^{\frac{7}{2}}S + k(k^2 S^2 + 1)^{\frac{1}{4}}(\sqrt{S + k^2} + 1))k^{\frac{5}{2}}S + \sqrt{2}k^3 S \\
&\leq (2k^{\frac{9}{2}} + 2k^{\frac{7}{2}} + \sqrt{2}k^3 + k)S + 2(k^5 + k^4 + k^{\frac{7}{2}})S^{\frac{3}{2}} + 2(k^6 + k^4)S^2 \\
&\leq 7k^{\frac{9}{2}}S + 6k^5 S^{\frac{3}{2}} + 4k^6 S^2.
\end{aligned}$$

Now $17k^6 S$ is a bound of the right hand side if $S \leq 1$. Otherwise, it is also an upper bound for the Kolmogorov distance since the Kolmogorov distance is always less or equal than one. Hence, (22) holds for U-statistics of the form (24).

A U-statistic $F \in L^2(\mathbb{P})$ with an arbitrary function $f \in L_s^1(X^k)$ can be approximated by a sequence $(f^{(n)})_{n \in \mathbb{N}}$ such that $|f_n| \leq |f|$ μ^k -almost everywhere, $f^{(n)}$ is bounded and $\mu^k(\text{supp}(f^{(n)})) < \infty$ for all $n \in \mathbb{N}$. As in [10], we define U-statistics

$$F^{(n)} = \sum_{(x_1, \dots, x_k) \in \eta_{\neq}^k} f^{(n)}(x_1, \dots, x_k),$$

that converge to F almost surely, in $L^1(\mathbb{P})$ and in $L^2(\mathbb{P})$. Using the triangle inequality for the Kolmogorov distance, we have

$$d_K \left(\frac{F - \mathbb{E}F}{\sqrt{\text{Var } F}}, N \right) \leq d_K \left(\frac{F - \mathbb{E}F}{\sqrt{\text{Var } F}}, \frac{F^{(n)} - \mathbb{E}F^{(n)}}{\sqrt{\text{Var } F^{(n)}}} \right) + d_K \left(\frac{F^{(n)} - \mathbb{E}F^{(n)}}{\sqrt{\text{Var } F^{(n)}}}, N \right).$$

In the first summand on the right hand side, we have almost sure convergence so that it vanishes as $n \rightarrow \infty$. For the second part, we can apply the bound (22) and replace $f^{(n)}$ by f since $M_{ij}(f^{(n)}) \leq M_{ij}(f)$ for all $i, j = 1, \dots, k$. \square

In a similar way, one can also obtain an upper bound for the Kolmogorov distance between a Gaussian random variable and a finite sum of Poisson U-statistics. This class of Poisson functionals is interesting since Theorem 3.6 in [10] tells us that each Poisson functional $F \in L^2(\mathbb{P})$ of order k with kernels $f_i \in L_s^1(X^i) \cap L_s^2(X^i)$ for $i = 1, \dots, k$ is a finite sum of Poisson U-statistics (and a constant). In the proof of Theorem 4.2, we derive upper bounds depending on R_{ij} and \tilde{R}_i for some of the terms occurring in (3). These bounds still hold for a Poisson functional $F \in L^2(\mathbb{P})$ that is a finite sum of Poisson U-statistics. Moreover, we can compute a similar bound as in Lemma 5.3 using the representation of F as a sum of Poisson U-statistics. Together with the fourth centered moment of F , it gives us an upper bound for the Kolmogorov distance between F and a Gaussian random variable.

Acknowledgement: This paper was written during a stay at Case Western Reserve University (February to July 2012) supported by the German Academic Exchange Service.

References

- [1] CHEN, L.; GOLDSTEIN, L.; SHAO, Q. (2011): *Normal Approximation by Stein's Method*, Springer, Berlin.
- [2] CHEN, L.; SHAO, Q. (2005): *Stein's Method for normal approximation*, in A. Barbour; L. Chen: An Introduction to Stein's Method, Singapore University Press, Singapore.
- [3] LACHIÈZE-REY, R.; PECCATI, G. (2011): *Fine Gaussian fluctuations on the Poisson space, I: contractions, cumulants and geometric random graphs*, arXiv:1111.7312v1 [math.PR].
- [4] LACHIÈZE-REY, R.; PECCATI, G. (2011): *Fine Gaussian fluctuations on the Poisson space, II: rescaled kernels, marked processes and geometric U-statistics*, arXiv:1205.0632v1 [math.PR].
- [5] LAST, G.; PENROSE, M.D. (2011): *Poisson process Fock space representation, chaos expansion and covariance inequalities*, Probab. Th. Rel. Fields. **150**, 663–690.
- [6] LAST, G.; PENROSE, M.D.; SCHULTE, M.; THÄLE, C. (2012): *Moments and central limit theorems for some multivariate Poisson functionals*, arXiv:1205.3033v1 [math.PR].
- [7] NUALART, D.; VIVES, J. (1990). *Anticipative calculus for the Poisson process based on the Fock space*, Lecture Notes in Math. **1426**, 154–165.
- [8] PECCATI, G.; SOLÉ, J.L.; TAQQU, M.S.; UTZET, F. (2010): *Stein's method and normal approximation of Poisson functionals*, Ann. Probab. **38**, 443–478.

- [9] PECCATI, G.; TAQQU, M.S. (2011): *Wiener Chaos: Moments, Cumulants and Diagram Formulae: A survey with computer implementation*, Springer, Berlin.
- [10] REITZNER, M.; SCHULTE, M. (2011): *Central limit theorems for U-statistics of Poisson point processes.*, arXiv: 1104.1039v2 [math.PR].
- [11] STEIN, C. (1972): *A bound for the error in the normal approximation to the distribution of a sum of dependent random variables*, in: Proceedings of the Sixth Berkley Symposium on Mathematical Statistics Probability 2, Univ. California Press, Berkley.
- [12] STEIN, C. (1986): *Approximate Computation of Expectations*. Institute of Mathematical Statistics, Hayward, CA.
- [13] SURGAILIS, D. (1984): *On multiple Poisson stochastic integrals and associated Markov semigroups*, Probab. Math. Statist. **38**, 217–239.